1993 (274)

Social Neuroscience Bulletin, Vol. 6, No. 2, Spring 1993, pp. 15-19.

A Mentalist View of Consciousness*

Roger W. Sperry
California Institute of Technology

The position that I support as "mentalism," "emergent mentalism," or just the "new mentalism" has been variously interpreted to be "dualism," a "mind-brain identity" view, "functionalism," "physical monism," "mental monism," "emergent materialism," and "dualistic interactionism"--among other things. Such extraordinary confusion may reflect a lack of clarity in my thinking and/or writing or, in part, my attempt to avoid specialized terminology in initial presentations to a broad general audience. Primarily, however, I believe it can be attributed to the nature of the mind-brain position itself, as pointed out by Natsoulas (1987) and more emphatically by Ripley (1984). The position is one which departs from previously accepted conceptual dichotomies in philosophy. As Natsoulas (1987, p. 1) states, it "blends together features of opposed solutions."

What follows is an effort to update and better clarify the theory by bringing together in a single listing some of the more contested and troublesome defining features as these have surfaced over time, along with a few comparisons to related contemporary views.

It will be helpful to make clear in advance that the following arguments for conceiving consciousness as a causal emergent in brain function are taken to apply as well to the new mentalist paradigm that recently has replaced behaviorist doctrine in psychology. The abrupt mentalist overthrow of the half-century old, seemingly impregnable behavioral paradigm (Baars 1986; Gardner 1985; Sperry 1987) with its secure basis in foundational tenets consistent with the other sciences, is hardly likely to have been achieved within a few years on different grounds via two dissimilar mentalistic theories.

As a result of the consciousness revolution of the 1970s in mainstream psychology (Dember 1974; Matson 1971; Palermo 1971; Pylyshyn 1973) there has emerged a paradigm conflict between different disciplines over what to use as the basic working conceptual framework for science. At stake is the question of what is valid in causal explanation.

On the one side, behavioral and cognitive science lead the way for a new *macromental* paradigm affirming downward (top-down) causal control by irreducible emergent properties over their component parts, a special instance of which is the downward control by mental events over the lower neuronal events. On the other side the more exact basic sciences continue predominantly to adhere to traditional microdeterministic, exclusive "bottom up" determination of the whole by the parts, in which the neuronal events determine the mental but not vice versa.

Essentially it translates into a conflict in the history of science over the kinds of forces claimed to control ourselves and the world. While the outcome should little affect the methodological approach or the actual day-to-day practice of science, it could vastly transform the kind of worldview science leads us to believe in, the relationship of science to values, and the type of causal explanation used in science and elsewhere.

The New Mentalism

> Subjective meaning depends on the over-all functional effects of the physiological processes, not upon their copying or representing in code form the attributes of stimlus (Sperry 1952, p. 308).
>
> The central excitation may vary considerably in its geometric spatiotemporal and other properties while maintaining invariant or equivalent functional value...[such that] the same psychic mechanisms may be obtained from brain patterns the neuronal details of which differ considerably on different occasions. Not only may different neurons be involved, as many configurationists (e.g., Gestaltists) would agree, but more than this, the configuration of neurons may vary (p. 309).
>
> Significance and meaning in brain function do not derive from the intrinsic protoplasmic or other analytic aspects of neural excitation, but rather from their higher-order functional and operational effects as these work upon successive brain states.... We should not expect to find that a single neuron or an isolated patch of neurons, or even a cortical center, could sense, feel, experience, or think anything in isolation (pp. 310-311).

This "functional" approach has been developed extensively in recent years by Foder (1981) and others, to become the central philosophy of the new computer oriented cognitive science and is thought of as a philosophic development of the late twentieth century (Boden 1988). In the above early statement, it was not conceived in an abstract restrictive computational or linguistic sense, but in more generalized terms that apply also to consciousness and emotion in the still-to-be elucidated codes of brain processing.

---

*Direct correspondence to Roger W. Sperry, Division of Biology 156-29, California Institute of Technology, Pasadena, California 91125, USA (fax: 1-818-449-0756).

Editor's excerpt from "Turnabout on Consciousness: A Mentalist View," *The Journal of Mind and Behavior*," Volume 13, Summer 1992, pp. 259-280. Republished with the permission of the *JMB* Editor.

Monism, Not Dualism

In responding to claims that the view I present is dualistic (Bindra 1970; Bunge 1959; Puccetti 1977), I explained at some length, under the titles "Mentalist Monism" (1978) and "Mentalism, Yes; Dualism, No" (1980), why this theory is not dualistic in the classical sense, but rather is a new form of monism that conceives mental entities in an emergent and causal role. In this view conscious experience, taken to be a nonreductive dynamic emergent of brain activity, cannot exist apart from the brain. There is no provision for unembodied consciousness, mind, or spirit. The early statement seems explicit.

> The present scheme would...eliminate the old dualistic confusions, the dichotomies and the paradoxes, proposing instead a single unified system... (Sperry 1965).

> [T]his represents a return toward mentalism except that the mental forces are not viewed in any metaphysical, preternatural, nonmaterial, epiphenomenal, or other dualistic sense (1970a, p. 588).

> The term [mental forces] fits the phenomena of subjective experience but does not imply here any disembodied supernatural forces independent of the brain mechanism. The mental forces as here conceived are inseparably tied to the cerebral structure and its functional organization (1970b, p. 137).

Mental States: Long-Banned, Now Made Legitimate for Science

The accepted basic assumption of materialist science that "mind does not move matter" (Herrick 1956, p. 281), that "no physical action waits on any thing but another physical action" (MacKay 1966, p. 438), or in neuroscience, that "As neurophysiologists we simply have no use for consciousness...to explain how the nervous system works" (Eccles 1966, p. 250) was contradicted and reversed. The traditional logic of causal explanation which we had previously assumed to be airtight and irrefutable was discovered to have an oversight or loophole allowing for a different form of causal determinism by which mental properties could be seen to have a supervenient form of objective causal control from above downward over the lower level neuronal events. Unlike previous approaches that had tried to insert the conscious effects into the causal matrix already described in neurophysiology, a new approach was used that preserved the microphysiology but embedded it within a (yet to be described) higher-level system that involves causal interaction of cognitive processes. Thus,

> ...the causal potency of an idea, or an ideal, becomes just as real as that of a molecule, a cell, or a nerve impulse (Sperry 1965, p. 82).

> ...the interplay of psychic and mental forces, though accessible like the interior of the earth only indirectly at this date becomes, in principle, a proper phenomenon for scientific investigation (p. 83). The present scheme would put mind back into the brain of objective science and in a position of top command (p. 85).

Brain and Mind

In contrast to mind-brain identity theory, in the present theory the conventional differences between mind and brain are accepted and emphasized. The brain process consists of a multilevel hierarchic compound of entities and events from subatomic up through cerebral circuit and cognitive levels, is composed mostly of nonconscious elements that are common both to conscious and unconscious brain processes. Only upper levels within certain brain processes are presumed to sustain consciousness qualities. The brain process and the conscious properties are inseparable but different. The difference between mind and brain is the kind which exists between an emergent property and its infrastructure.

> Subjective experience in this interpretation is conceived to be an emergent dynamic property of cerebral excitation, inseparable from the material bain process, but different from and more than, the collected sum of the physicochemical components (Sperry 1969).

Emergent mental images and percepts typically have a subjectively experienced unity, continuity, and constancy markedly unlike the disparate, scattered and often spatially splintered and transformed neuronal events that generate the conscious experience. Early illustrations were given in reference to the neural events involved, for example, in perceiving a simple triangle.

> This approach does not lead us to expect in the cerebral process any kind of triangularity, linearity, nor even a unity corresponding to that of the perceived triangle.... When the fixation point shifts to either side, the fragments of the triangle projected to each hemisphere change accordingly. As the eyes rove over the triangle from apex to base and from side to side, the shape and also the position of the cortical patterns change radically in each hemisphere

> While the kaleidoscopic series of excitatory changes is taking place in the visual cortex of the brain, the figure of subjective experience remains constant, a unified whole, with a fixed orientation in space.... [I]t is contended that unity in subjective experience does not derive from any kind of parallel unity in the brain processes. Conscious unity is conceived rather as a functional or operational derivative...with nowhere a compact unified pattern of discharge that represents "triangle" (Sperry 1952, pp. 300-307).

A similar argument was later used by Eccles (Popper and Eccles 1977) as a key component of his hypothesis for psychoneural interaction. In the present view, however, the difference between mental and neuronal is no greater than the difference between emergent properties and their components. The conscious experience is a property of, and cannot exist apart from, the cerebral substrate.

Subjective Conscious Qualities: A Causal Control Role

The central distinguishing tenet of the new materialism is contradictory to the centuries-old tenet of science that mind does not move matter. It is in direct conflict as well with behaviorism's decades-old denunciation of consciousness and of mental explanations.

> ...I find it difficult to believe that the sensations and other subjective experiences per se serve no function, have no operational value and no place in our working models of the brain.... [P]ain per se, and subjective experience in general, emerged in central nervous evolution and could only have been maintained and differentiated because it does serve a real use, i.e., by virtue of its operational value in the causal sequence. On these terms one wonders if any physiological model...that fails to include the subjective properties is not bound to end up with some kind of gap in the chain of cerebral events (Sperry 1959, pp. 420-421).

That conscious qualities themselves as subjectively experienced can exert causal control or influence on the course of physical events in the brain still remained a brash, unacceptable concept from the standpoint of science when I first tried to spell it out more fully as "An Alternative Mentalist Position" (Sperry 1965, p. 78).

> ...it comes down to the issue of who pushes whom around in the population of causal forces that occur in the cranium... of straightening out the peck-order hierarchy among intracranial control agents. Even the brain cells...do not have very much to say about when they are going to fire their message, for example, or in what time pattern.... The firing orders for the day comefrom a higher command. In other words, the flow and the timing of impulse traffic through any brain cell, or even a nucleus of cells in the brain are governed largely by the over-all encompassing properties of the whole cerebral circuit system (Sperry1965, p. 79).

> The reference here is not to the whole circuit system of the entire brain, but only to that of the particular cerebral process for sustaining a given mental state or experience.

> [I]f one keeps climbing upward in the chain of command within the brain, one finds at the very top those over-all organizational forces and dynamic properties of the large patterns of cerebral excitation that are correlated with mental states or psychic activity (Sperry 1965, p. 80)

> In the onward flow of conscious brain states, one state calling up the next, these are the kinds of dynamic entities that call the plays. It is exactly these encompassing mental forces that direct and govern the inner impulse traffic, including its electrochemical and biophysical aspects (p. 82).

Top-Down Control

Referred to also as emergent determinism, molar, holistic, or macro-determinism, downward causation, emergent causation and, beginning in the 1970s, as "systems" thinking, a "system's" view--this revised concept of causal control is a critical key feature of the new mentalist paradigm, and often the feature that causes the most misunderstanding. This is the feature also that stands most directly in contradiction with the orthodox bottom-up form of causality previously relied on in traditional materialism. In debates about the consciousness revolution we thus tend to focus eventually on the nature and role of downward causal determination.

> . . .a molecule in many respects is the master of its inner atoms and electrons. The latter are hauled and forced about in chemical interactions by the overall configurational properties of the molecule as a whole (Sperry 1964, p. 2).

But we have always been taught the converse, that things work the other way around, i.e., that the inner atoms of a molecule determine its properties and behavior in chemical reactions. This latter still continues to be the doctrine most commonly adhered to in present-day physics and chemistry. Our new view does not deny this but holds that in addition there are reciprocal emergent forces of the molecule as a whole that also exert concomitant control from above downward over the parts embedded at different levels within the structural hierarchy of the molecule.

Many physicists of the old persuasion protest that one can predict emergent properties such as the wetness of water from knowledge of the hydrogen and oxygen atoms alone in accord with the standard view that everything is determined below upward following the course of evolution, and therefore that additional "emergent properties and forces" are not needed. This classic microdeterministic view of evolution and physical reality gets replaced in the new macro-mental reasoning by a different, upgraded picture:

> . . . evolution keeps complicating the universe by adding new phenomena that have new properties and new forces and that are regulated by new scientific principles and new scientific laws.... [T]he old simple laws...never get lost or canceled in the process of compounding the compounds. They do, however, get superseded, over-whelmed, and outclassed by the higher-level forces as these successively appear... (Sperry 1964, p. 2).

> At the top, in the human brain, these include the powers of perception, cognition, reason, judgment and the like, the operational causal effects and forces of which are equally or more potent in brain dynamics that are the outclassed inner chemical forces (p. 20).

Downward determinism gained a strengthened scientific status in the early 1970s when mainstream psychology underwent a shift from behaviorism to mentalism, a shift for which downward determinism (of the mental over the neuronal) is a logical necessity. The cause for this turnaround by a whole scientific discipline (the "consciousness/cognitive revolution") is still in dispute, but clearly involves a large complex of factors: sociological, intuitive and wishful, as well as scientific and rational (Baars 1986; Gardner 1985; Sperry 1987). Underneath it all, however, the new paradigm had to contain a challenging core concept that was both sound and also incompatible with prevailing behaviorist doctrine such that the behaviorist paradigm could not be stretched to encompass it (Cohen 1985; Kuhn 1970; Sperry 1987). Otherwise, fervent adherents of behaviorism would quickly ferret out any vulnerability in the new concept or show behaviorism to be adequate to include it.

Top-down control of the parts by the emergent whole is such a core concept; it legitimizes consciousness and the subjective for science. Introducing a modified basis for causal explanation, it stands in direct conflict with the conventional exclusive microdeterministic premises that exclude consciousness and mental explanation--not only in behavioral science but also in neuroscience and in all basic science. This was the first time that a counter conception had been found that could successfully refute the essence of behaviorism per se as a paradigmatic worldview and philosophy of science (Reese and Overton 1972; Skinner 1964). The new concept contradicts as well the microdeterminist paradigm of physical science in general.

A Unifying Middle-Way View

For understanding the mind-brain relation and associated problems, we had a choice in the past between mutually opposed explanatory frameworks: monism vs. dualism, materialism vs. mentalism, positivism vs. epiphenomenalism or existentialism, and, in the "two cultures" of C.P. Snow: the value-free deterministic materialist description of science versus the traditional humanistic (value laden, purposeful) views of the rest of society (Jones 1965).

The two formerly conflicting and mutually exclusive approaches become reconciled in the present view within a single monistic hierarchy by treating mental states as non-reducible emergent properties of brain processing that exert sequential and downward causal influence, but have no independent existence of their own part from the brain:

> ...note that the earlier distinction or dichotomy between mentalism and materialism is resolved in this interpretation...proposing instead a single unified system extending from subnuclear forces near the bottom up through ideas at the top. As a scientific theory of mind, it would provide a long sought unifying view on which to base our conceptions of human nature... (Sperry 1965, pp. 84 85).

Emergent Macro-Processes Throughout Nature Become Irreducible and Ineliminable Causal Constructs

Emergent/holist theory, following an initial strong upsurge in the 1920s and 1930s, underwent a tapering off and decline during the 1940s. By the late 1950s it had fallen into considerable disfavor in philosophy, psychology, and biology (Feigl 1967; Köhler 1960). After the mid 1960s emergent theory burst into another upsurge of new interest (Koestler and Smythies 1969; Bertalanffy 1968; Polanyi 1968a, 1968b) that still is strongly sustained.

The new approach shifted the focus from how the emergent properties are formed to the causal consequences of their formation, making a major difference with respect to scientific explanation. In the revised outlook the emergent properties are causally determined.

> ...I hold that every time the elements of creation, whether atoms or concepts, are put together in the same way under the same conditions, that the same new properties would emerge and that the emergent process is, there fore, causal and deterministic. To this extent and in this sense it may also be said to be, in principle, predictable though generally, with few expectations, it is not so in practice (Sperry 1980, p. 200).

That emergent properties typically tend to be dramatically novel and rarely predictable is not because of any mystical unembodied forces (Smart 1981) but rather because of the qualitative complexity of the causal factors, arising particularly in the spacing and timing relations of the parts, plus the lack as yet of scientific principles for the types of causality involved. The interlevel upward and downward doubly determinate, or reciprocal form of causation involves two very different types of causal determination, both of which are in process simultaneously and without any conflict.

The "macromental" paradigm is actually a micro-macromental model of causal determination in which the microdeterminism is taken for granted and emphasis placed on the new macro and mental features. The mental, though only a subset of the macro, is sufficiently special in its importance and impact as to merit specific inclusion in the terminology.

The same reasoning by which conscious mental states become causal applies as well to unconscious mental states. The same holistic principles apply universally to hierarchically organized entities throughout nature at all levels: molecular, cellular, organismic, social and higher, and to inanimate as well as animate systems. The reasoning affirms the autonomy of the various different sciences at their own level and on the same grounds reinstates Newtonian physics as a realm of science not overthrown by, nor subsumed under, quantum mechanics. The same changed concept of causality negates the ideabehind the "unity of the sciences" movement and also that behind a "theory of everything" discoverable at some ultimatesubatomic level.

REFERENCES

Baars, R.J. (1986) *The Cognitive Revolution in Psychology.* New York: Guilford

Bertallanfy, L. von (1968) *General Systems Theory.* New York: Braziller.

Bindra, D. (1970) "The Problem of Subjective Experience: Puzlment on Reading R. W. Sperry's "A Modified Concept of Consciousness"." *Psychological Review* 77:581-584

Boden, M.A. (1988) *Computer Models of Mind.* NewYork: Cambridge University Press

Bunge, M. (1959) *Causality.* Cambridge: Harvard Univesity press

Cohen, B. (1985) *Revolution in Science.* Cambridge: Belknap Press of Harvard University Press.

Dember, W.N. (1974) "Motivation and the Cognitive Revolution. *American Psychologist* 29:161-168.

Eccles, J.C., ed. (1966) *Brain and Conscious Experience.* NewYork: Springer

Feigl, H. (1967) *The "Mental" and the "Physical. ."* Minneapolis: University of Minnesota Press

Fodor, J.A. (1981) "The Mind-Body Problem." *Scientific American* 244:114-123

Gardner, H. (1985) *The Mind's New Science: A History of the Cognitive Revolution.* New York: Basic Books.

Herrick, C.J.(1956) *The Evolution of Human Nature.* Austin: University of Texas Press.

Jones, W.T. (1965) *The Sciences and the Humanities.* Berkeley: University of California Press.

Koestler, A. and J.R. Smythies, eds. (1969) *Beyond Reductionism: New Perspectives in the Life Sciences.* (The Alpback Symposium 1968). London: Hutchinson.

Köhler, W. (1960) "The Mind-Body Problem." Pp. 14-32 in S. Hook (ed.), *Dimensions of Mind.* New York: Collier Press.

Kuhn, T. (1970) *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press.

MacKay, D.M. (1966) "Cerebral Organization and the Conscious Control of Action." Pp. 422-455 in J.C. Eccles (ed.), *Brain and Conscious Experience.* New York: Springer.

Matson, F.W. (1971) "Humanistic Theory: The Third Revolution in Psychology." *The Humanist* 31(2):7-11.

Natsoulas, T. (1987) "Roger Sperry's Monist Interactionism." *The Journal of Mind and Behavior* 8:1-12.

Palermo, D.S. (1971) "Is A Scientific Revolution Taking Place." *Science Studies* 1:135-155.

Polanyi, M. (1968a) "Life's Irreducible Structure." *Science* 160:1308-1312.

——— (1968b) "Logic and Psychology." *American Psychologist* 23:27-43.

Puccetti, R. (1977) "Sperry on Consciousness: A Critical Appreciation. "*Journal of Medicine and Philosophy*2:127-144

Popper, K.R. and J.C. Eccles (1977) *The Self and Its Brain.* New York: Springer International.

Pylyshyn, Z.W. (973) "What the Mind's Eye Tells the Mind's Brain:A Critique of Mental Imagery." *Psychological Bulletin* 80:1-24.

Reese, H.W. and W.F. Overton (1972) "On Paradigm Shifts." *American Psychologist* 27:1197-1199

Ripley, C. (1984) "Sperry's Concept of Consciousness." *Inquiry* 27:399-423.

Skinner, B.F. (1964) "Behaviorism at 50." Pp. 79-108 in T. Wann (ed.), *Behaviorism and Phenomenology* . Chicago: University ofChicago Press.

Smart, J.J.C. (1981) "Physicalism and Emergence." *Neuroscence* 6:109-113.

Sperry, R.W. (1952) "Neurology and the Mind-Brain Problem." *American Scientist* 40:291-312.

——— (1959) "Discussion." Pp. 420-421 in M.A.B. Brazier (ed.), *The Central Nervous System and Behavior.* Princton, NJ: Madison Print.

——— (1964) "Problems Outstanding in the Evolution of Brain Function." *James Arthur Lecture on the Evolution of the Human Brain.* New York: American Museum of Natural History.

——— (1965) "Mind, Brain, and Humanist Values." Pp. 71-92 in J.R. Platt (ed.), *New Views of the Nature* of Man. Chicago:University of Chicago Press.

——— (1969) "Toward a Theory of Mind." [Abstract] *Proceedings of the National Academy of Sciences* 63:230-231.

——— (1970a) "An Objective Approach to Subjective Experience: Further Explanation of a Hypothesis." *Psychological Review* 77:585-590.

——— (1970b) "Perception in the Absence of Neocortical Commissures." *Research Publications of the Association for Research in Nervous and Mental Disease* 48:123-138.

——— (1978) "Mentalist Monism: Consciousness as a Causal Emergent of Brain Processes." *Behavioral* and Brain Sciences 3:365-367.

——— (1980) "Mind-Brain Interaction: Mentalism, Yes: Dualism, No." *Neuroscience* 5(2):195-206.

——— (1987) "Structure and Significance of the Consciousness Revolution." *The Journal of Mind and Behavior* 8:37-66.